Feature Selection for Robust Color Image Retrieval

Madirakshi Das and Edward M. Riseman[*]
Multimedia Indexing and Retrieval Group
Joint group of Center for Intelligent Information Retrieval and Computer Vision laboratories
Department of Computer Science
University of Massachusetts, Amherst, MA 01003-4610
E-MAIL: mdas,riseman@cs.umass.edu
HOMEPAGE: http://vis-www.cs.umass.edu/~mdas

Abstract

This work addresses the issue of color feature selection for content-based retrieval from large, heterogeneous color image databases where no assumptions can be made about the images or the type of queries. The color features used to describe an image have been developed based on the need for speed in matching and ease of computation on complex images while maintaining invariance to differences in scale, orientation, and location of the queried object in the database images and also the presence of significant, interfering backgrounds. The colors present and their spatial relationships are used as features to describe a color image. These features are used in an efficient, multi-phase retrieval system to produce retrieval results fast enough for use with an online user. Test results with multi-colored query objects from man-made and natural domains highlight the capabilities of the system.

1 Introduction

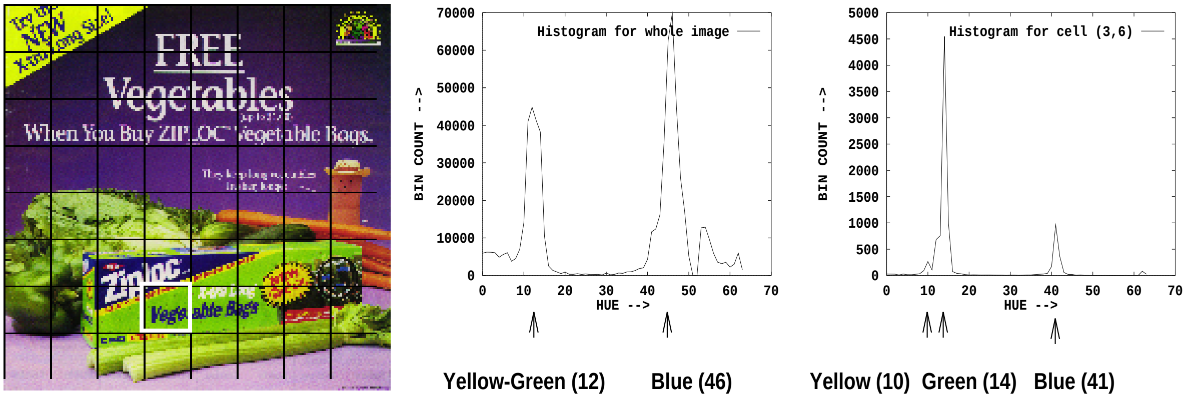
With the growing number of multimedia databases, retrieval of images, audio and video from large databases has become an active area of research. As in text retrieval, objects in the database which are relevant to the query being posed by the user need to be retrieved. Content-based retrieval is a popular paradigm for ensuring relevance in image retrieval where the aim is to find the images in a database which contain the object represented in a query image.

The fact that there are no obvious features across images which carry semantic information like words do in text, makes the selection of descriptive features for an image difficult. However, when the database has images of multi-colored objects which can be recognized on the basis of their distinctive color signatures, the color of the object and related color-based features are an obvious choice for indexing.

There has been work in color-based retrieval using color histograms [Swain and Ballard, 1991][Hafner *et al.*, 1995], but the retrieval results are sensitive to difference in scale and viewpoint between the object as depicted in the query image and as present in the database images. Using color clusters [Kankanhalli *et al.*, 1996] avoids the scale problem but both strategies are affected by the presence of interfering backgrounds, particularly when the query image is a small embedded part in a large target image. We have considered the general case where multi-colored objects occur with sig-

---

[*] This material is based on work supported in part by the National Science Foundation, Library of Congress and Department of Commerce under cooperative agreement number EEC-9209623, in part by the United States Patent and Trademark Office and Defense Advanced Research Projects Agency/ITO under ARPA order number D468, issued by ESC/AXS contract number F19628-95-C-0235, in part by the National Science Foundation, Central Intelligence Agency, Department of Defense (DARPA) and National Security Agency under grant number IRI-9619117, in part by ARPA (via USAF Rome Laboratory) under contract F30602-94-C-0042 and in part by NSF Multimedia CDA-9502639. Any opinions, findings and conclusions or recommendations expressed in this material are the authors and do not necessarily reflect those of the sponsors.



(a) (b) (c)

Figure 1: Effect of interfering background on histogram peak location: (a) "Ziploc" advertisement with a cell highlighted (b) global hue histogram of whole image with relevant peaks labelled (c) hue histogram of marked cell with peaks labelled

nificant, interfering backgrounds and in widely varying sizes, locations and orientation in the database images.

2 Selection of features

The main requirement for the color characteristics selected for matching is to provide discrimination between images which contain objects similar to the query object and those which do not. The feature matched needs to be invariant to differences in the scale, location and orientation of the query object in the candidate image and the presence of background colors in the candidate image. It is also desirable for the characteristics to be indexable and the matching process to be fast.

We have used two scale and orientation invariant color features, describing the *color content* of an image and the *spatial relationships* between color regions. In general, there is a trade-off between the discriminatory power of a color feature and its speed of matching. Simpler features are easy for indexing and matching, while complex features which provide more discriminatory power may not be indexable and take longer to match. We have selected both types of features and employed a two phase matching strategy to balance the trade-off between speed of retrieval and the precision obtained. The emphasis in the first phase of matching is on speed of retrieval, and the second phase aims at removal of false matches from the image list produced by the first phase.

2.1 Histogram Peaks as features

The simplest constraint on a database image retrieved as a response to a query is that it must have all the colors of the query object present. To check for this requirement, we need to describe the *color content* of an image. As observed in [Matas *et al.*, 1995], the *locations* of peaks in a histogram are stable under viewpoint change and scale transformation, unlike histogram bin *counts* used in [Swain and Ballard, 1991]. The storage space required is reduced when compared to using the full histogram, and standard key-based indexing techniques can be used. Even the peak locations are affected by the presence of background in the image. However, the color peaks present in an image can be determined more accurately when the histogram covers a small area of the image, minimizing the the presence of interfering colors from the background as illustrated in Figure 1.

We use a *split and merge* technique for peak detection which produces accurate peaks in spite of the presence of interfering backgounds. Since we do not know a priori the size or the location of the object of interest in the image, the image is divided uniformly into m x n cells. Local histograms are constructed for each image cell in the HSV (hue, saturation, value) color space since it is more stable than RGB under variations in illumination. Since the hue component is the most stable and value component the least stable, we use HSV histograms with finely discretized hue axis and coarsely quantized saturation and value axes (64x10x10). Peaks are detected in the histograms by finding local maxima in a 3-D neighborhood window. A combined list of peaks is produced by merging multiple copies of the same peak.

2.2 Describing spatial relationships between colors

There could be many images which have all the colors of the query object, but not in the same spatial configuration as in the query object. In the extreme case, the matched colors are scattered across the image and do not form any connected cluster. In other cases, some color adjacency relationship present in the query object may be violated in candidate images. For example, in the query in Figure 2 (a), the red (labelled 0) and blue (labelled 3) regions are adjacent whereas in the false match (b), they are not adjacent. These false matches could be eliminated if information on spatial distribution of colors in the image was available.

The color adjacency graph (CAG) formulation used by Kittler et al [Matas *et al.*, 1995] is a good descriptor of the color relationships in a multi-colored object, where the color regions are nodes in a graph with edges connecting color regions which share an edge at the pixel level. However, a CAG description of the database images is not feasible for retrieval due to the complexity of the images. Most of the images contain natural objects and color regions in which there are no distinct boundaries between colors. An attempt to construct a CAG for these images has produced very large, complex graphs, making the matching phase intractable. Therefore, we need a simpler representation for the spatial distribution of colors that allows efficient generation and storage for all images and allows fast matching.

(a) (b)

Figure 2: Example of mismatch in spatial color relationships : (a) “Blueberry Morning” query image (b) A false match

It should be noted that during the peak detection process, we have already localized color peaks in image cells, giving us the color content in each cell. We now use this information to construct a graph describing the *approximate* spatial relationships between colors in the image *without any additional pixel level processing.*

We start by constructing an intermediate graph representation directly from the peak description of the image based on whether pixel level adjacency is *possible* between two color regions, and condense it into a compact graph - the *spatial proximity graph* (SPG). Each node in the intermediate SPG corresponds to a detected color peak, and edges between two nodes indicate that the two color regions which produced the peaks could be adjacent in the image. Let nodes of the intermediate SPG be of the form c^i_m, where m is the peak color label of the node and i is the cell in which it is located. There is an edge E between two nodes of the graph if the following condition is met.

- $E(c^i_m, c^j_n)$ if $i = j$ OR $m = n$ and (i, j) are 4-neighbors.

The intermediate graph obtained is not scale invariant, since a larger region would produce more nodes in the graph. A smaller, scale invariant SPG which still captures the spatial relationships between colors is obtained by *collapsing* connected nodes of the same color label into a single node of that color label. The graph may still have multiple nodes of the same color label, but only if these peaks were spatially disconnected in the image. The SPG is computed off-line for all database images and stored using an adjacency matrix representation.

The spatial proximity graph (SPG) description has a number of very useful properties. Apart from being scale and orientation invariant, it can be computed easily for all types of images, with or without prominent color boundaries. The SPG shows all possible pixel-level adjacencies that could appear in an image, without going through pixel-level processing. So any color adjacency relationship present in the image is still captured in this simplified graph. On the other hand, the graph is approximate since it may indicate some possible adjacency relationships for which there is actually no pixel-level adjacency in the image e.g. when two color regions are within a cell in the image, but there is no pixel adjacency between them.

3 Overview of Retrieval System

The features desribed above are used in an experimental retrieval system, FOCUS (Fast Object-Color-based qUery System). The schematic diagram of the system is shown in Figure 3. When a query image is obtained, the peaks in its color histogram and the graph describing the color relationships are computed.

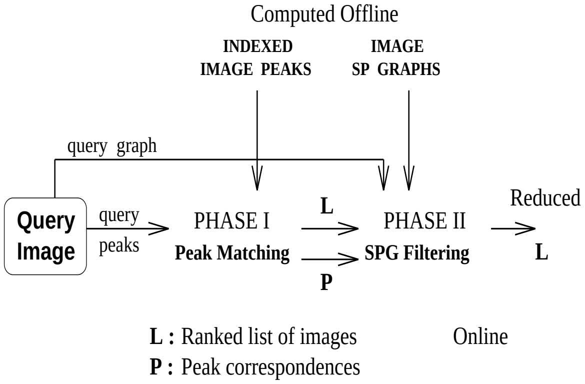
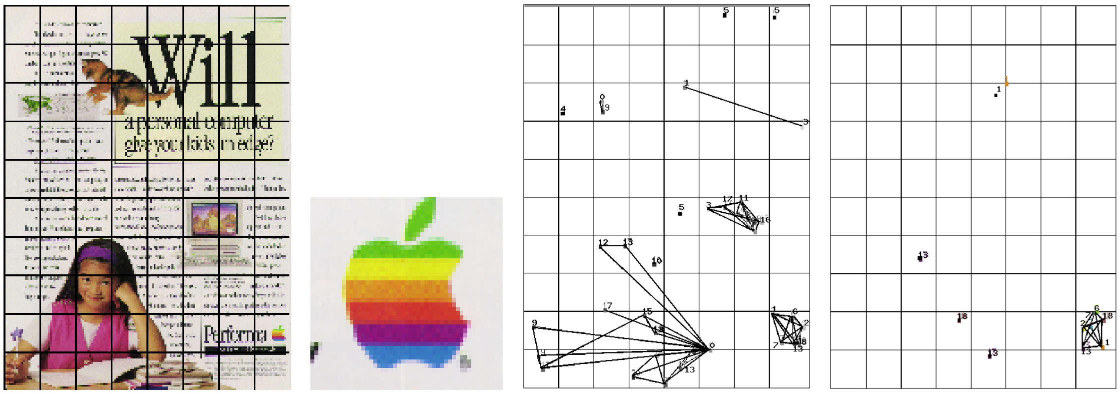


Figure 3: Overview of the FOCUS image retrieval system

The peaks extracted offline from the database images are stored in a $B+$ tree which is an order-preserving indexing structure. A *frequency table* is also constructed which gives the number of images which will be retrieved for each point in the discretized HSV space. For each peak in the query, P_q (h_q, s_q, v_q), a range query of $(h_q \pm 3, s_q \pm 4, v_q \pm 5)$ is executed starting with the peak which retrieves the minimum number of images onwards. A *join* of the lists of image identifiers is taken to find the images which have peaks matching *all* query peaks. The time complexity of the retrieval process is given by $O(qlog(kN))$, where q is the number of query peaks, N is the total number of images in the database and k is the average number of peaks per image. The images extracted are ordered by increasing mismatch scores, where the mismatch score is computed as the total *city block* distance between the matched candidate image peaks and the query peaks.

The correspondence between each color label in the image and the color peak in the query image which it matched is available from the peak matching computed during the first phase. Many image color labels may not match any query peak, since peaks maybe produced by the background in the image. The SPG computed off-line can be drastically reduced by removing all nodes in the image SPG whose color label does not match any query peak. The reduced SPG is also relabelled using the query peak color labels so that both the query graph and the reduced SPG now use the same color labels. The reduced SPG is much smaller than the original SPG, as illustrated in Figure 4, making the graph matching feasible as an online process.

The problem tackled during the online second phase is to detect if the query color graph occurs as a sub-graph of the reduced candidate image SPG. Though this is an instance of the subgraph isomorphism problem which is known to be NP-complete, due to the restricted nature of this problem, the matching computation is feasible. The running time is of the order of $O(n^m)$ where n is the size of the query adjacency matrix and m is the maximum number of instances of a color label in the reduced SPG, typically 3 or



(a) (b) (c) (d)

Figure 4: Reducing SPGs by deleting nodes not matched in phase 1: (a) Original image (b) Query image (c) SPG computed offline (d) SPG after reduction

less. Further details of this system can be found in [Das *et al.*, 1997][Das *et al.*, 1996].

4 Results

The database on which FOCUS has been tested consists of 400 advertisements from magazines and 800 color images from nature including birds, fish, flowers, animals land vegetables. The retrieval results obtained can be judged using *precision* and *recall* as criteria. Precision is the proportion of correct retrievals in the images retrieved upto the last correct image. Recall is the proportion of correct retrievals out of all the images in the database that should have been retrieved for the given On a query set of 25, the recall was 95%. The average precision after phase 1 was 44% and after phase 2 it improved to 60%. The performance was better when the query had more than three colors. The average precision score for a query set with more than three colors was 50% after phase 1 and 75% after phase 2. Two sample retrieval results are shown in Figure 5.

The time taken for a complete cycle of retrieval consists of the query processing time, phase 1 matching and phase 2 matching. FOCUS runs on a 133 MHz Pentium processor and all times mentioned are averaged over many trials. Query processing takes about 0.1 sec on a query image of size 100x200, which is the average size of queries tried. Phase 1 matching takes 0.1-0.2 sec and phase 2 matching takes about 0.01 for each image in the list produced by phase 1. Since this list has 30 images on an average the second phase takes about 0.3 sec. The retrieval process is fast enough to be scalable to very large databases since the query processing time is independent of the size of the database, the first phase of matching grows only logarithmically with the size of the database and the second phase depends only on the number of images retrieved by the first phase.

5 Conclusion

We have presented two robust color features which have been used to develop a fast, background independent color image retrieval system which produces good results with multicolored query objects. The retrieval is robust to differences in the scale, orientation and location of the query object in candidate images. The speed of the system and the small storage overhead make it suitable for use in large databases with online user interfaces. In future, we plan to increase the size of the database, add more color features to further distinguish between images and add more phases to utilize other types of image information.

Figure 5: Examples of Retrieved Results - the query is marked by a white box

Acknowledgments

We would like to thank R.Manmatha for testing the retrieval results and providing valuable suggestions, Jonathan Lim for building the webpage for the online demonstration of the system and Bruce Draper for ideas in earlier versions of this work.

References

[Swain and Ballard, 1991] M.J. Swain and D.H. Ballard. Color Indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.

[Niblack *et al.*, 1993] W. Niblack, R. Barber *et al.* The QBIC project: Querying Images by Content using Color, Texture and Shape. *SPIE Conference on Storage and Retrieval for Image and Video Databases*, 1908:173–187, 1993.

[Hafner *et al.*, 1995] J. Hafner, H. Sawhney, W. Niblack *et al.* Efficient Color Histogram Indexing for Quadratic Form Distance Functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(7):729-736, 1995.

[Matas *et al.*, 1995] J. Matas, R. Marik and J. Kittler. On Representation and Matching of Multi-Coloured Objects. *Fifth International Conference on Computer Vision*, 726-732, 1995.

[Kankanhalli *et al.*, 1996] M.S. Kankanhalli, B.M. Mehtre and J.K. Wu. Cluster-Based Color Matching for Image Retrieval. *Pattern Recognition*, 29(4):701-708, 1996.

[Gong *et al.*, 1996] Y. Gong, C.H. Chuan and G. Xiaoyi. Image Indexing and Retrieval Based on Color Histograms. *Multimedia Tools and Applications*, 2(2):133-156, 1996.

[Das *et al.*, 1996] M. Das, B.A. Draper, W.J. Lim, R. Manmatha and E.M. Riseman. A Fast, Background-independent Retrieval Strategy for Color Image Databases. *Computer Science Technical report* TR-96-79, Univ. of Massachusetts at Amherst.

[Das *et al.*, 1997] M. Das, E.M. Riseman and B.A. Draper. FOCUS : Searching for Multi-colored Objects in a Diverse Image Database. *To appear in IEEE conference on Computer Vision and Pattern Recognition*, 1997.